Stochastic Processes and their Applications 3 (1975) 45–54

ASYMPTOTIC NORMALITY OF THE NUMBER OF SMALL DISTANCES BETWEEN RANDOM POINTS IN A CUBE *

Adri KESTER
University of Rochester, Rochester, N.Y. 14627, U.S.A.

Received 10 December 1973
Revised 5 August 1974

Draw n random points from a cube, and consider the number of distances smaller than r between those points. When $n \to \infty$, $r \to 0$ simultaneously, and within certain relations to each other, this number is seen to be asymptotically normal.

asymptotic normality
random distances
moment convergence

0. Introduction

Let $E^k = [0, 1]^k$ be the unit cube in $\mathbf{R}^k$, the k-dimensional space of real numbers, and denote by $B(r)$ the ball in $\mathbf{R}^k$ with radius r and center 0. Draw n points $x_1, x_2, ..., x_n$ independently and at random from E^k, and write $t_n(B(r))$ for the number of pairs (i, j), $1 \leqslant i < j \leqslant n$, for which $x_i - x_j \in B(r)$. Alternatively, $t_n(B(r))$ is the number of distances between the n chosen points smaller than r. It is shown that $t_n(B(r))$ is asymptotically normally distributed if $r = r(n)$ tends to zero with a certain speed as $n \to \infty$ (Theorem 2.1). More generally, if $B_1, B_2, ..., B_m$ are disjoint subsets of $B(r)$, it is shown that under the same circumstances the vector $(t_n(B_1), t_n(B_2), ..., t_n(B_m))$ is asymptotically multivariate normal with independent components. The proofs of these theorems will use moment convergence and a few elementary notions from graph theory.

* An earlier version of this paper was written at the University of Amsterdam.

1. A homogeneous problem

As the problem is stated, there are boundary effects in that the distribution of the distance between a fixed point x and a random point y depends on the position of x in the cube. To eliminate these boundary effects, consider $x_1, x_2, ..., x_n$ modulo 1 in each component. In effect, one can imagine the x_i drawn randomly and independently from $C^k := ([0,1] \bmod 1)^k$. Let λ be the Lebesgue measure on C^k or E^k as the case may be. Since $\lambda(E^k) = \lambda(C^k) = 1$, the x_i are distributed according to λ. Working in C^k rather than E^k has the following advantage.

Lemma 1.1. *On C^k, the random vectors $x_1, x_2 - x_1, x_3 - x_1$, are independent and identically distributed.*

Proof. For each measurable set $A \subset C^k$, and $x \in C^k$, the conditional probability

$$\mathbf{P}[x_2 - x_1 \in A \mid x_1 = x] = \lambda(x + A) = \lambda(A)$$

is independent of x since λ is invariant under translation. Further,

$$\lambda(A) = \mathbf{P}[x_1 \in A],$$

so that x_1 and $x_2 - x_1$ are i.i.d. The other necessary conditions are checked in analogous ways. □

Now for any Borel set $B \subset C^k$, let $s_n(B)$ be the number of pairs (i, j) $1 \leqslant i < j \leqslant n$ such that $x_i - x_j \pmod 1)^k \in B$. Let $A \subset B(r) \subset \mathbf{R}^k$ for some $0 < r < \frac{1}{2}$, and let $A^* \subset C^k$ be given by

$$A^* = \{x \pmod 1)^k \mid x \in A\}.$$

Then $\lambda(A) = \lambda(A^*)$. Define the following binary variables:

$$d_{ij} = \begin{cases} 1 & \text{if } x_i - x_j \in A, \\ 0 & \text{otherwise} \end{cases} \tag{1.1a}$$

$$b_{ij} = \begin{cases} 1 & \text{if } x_i - x_j \pmod 1)^k \in A^*, \\ 0 & \text{otherwise.} \end{cases} \tag{1.1b}$$

Then

$$s_n = s_n(A) = \sum b_{ij}, \qquad t_n = t_n(A) = \sum d_{ij},$$

where both summations are over i, j with $1 \leqslant i < j \leqslant n$. The r.v.'s b_{ij} are pairwise independent (to see that b_{12} and b_{13} are independent, use Lemma 1.1), but not mutually independent (e.g. the triple b_{12}, b_{13}, b_{23} is not independent). Writing λ for $\lambda(A)$, one finds from Lemma 1.1 that

$$\mathrm{E}\{b_{ij}\} = \mathrm{P}[x_i \in A^*] = \lambda, \qquad \operatorname{var} b_{ij} = \lambda(1-\lambda),$$

so that

$$\mathrm{E}\{s_n\} = \tfrac{1}{2}n(n-1), \qquad \operatorname{var} s_n = \tfrac{1}{2}n(n-1)\,\lambda(1-\lambda). \tag{1.2}$$

In calculating var s_n, Lemma 1.1 is used to see that the covariances vanish.

To compare s_n with t_n, an asymptotic result is needed.

Theorem 1.2. *If A varies with r in such a way that $A \subset B(r)$ as $r \downarrow 0$, then for $n \to \infty$ and $r \downarrow 0$ one has*

$$\frac{\operatorname{var}(s_n - t_n)}{\operatorname{var} s_n} = \mathrm{O}(r) + \mathrm{O}(nr\lambda), \tag{1.3}$$

where $\lambda = \lambda(A)$.

Proof. Put

$$c_{ij} = b_{ij} - d_{ij}.$$

Since $x_i - x_j \in A$ implies $x_i - x_j \in A^*$, c_{ij} takes only the values 0,1. Consider

$$\operatorname{var}(s_n - t_n) = \operatorname{var} \sum c_{ij}.$$

Obviously, $\operatorname{cov}(c_{ij}, c_{kl}) = 0$ whenever i, j, k, l are all different. Since there are $n(n-1)(n-2)$ pairs (c_{ij}, c_{kl}) with exactly one suffix value occurring twice, one has

$$\operatorname{var}(s_n - t_n) = \tfrac{1}{2}n(n-1) \operatorname{var} c_{12} + n(n-1)(n-2) \operatorname{cov}(c_{12}, c_{13}).$$

From (1.2) it follows that one need only show for $r \downarrow 0$ and $A \subset B(r)$,

(a) var $c_{12} = \mathrm{O}(r\lambda)$,

(b) $\operatorname{cov}(c_{12}, c_{13}) = \mathrm{O}(r\lambda^2)$.

Let $D = E^k - [r, 1-r]^k$ be a layer of width r on the inside of the boundary of E^k. From (1.1) it can be seen that $c_{12} = 1$ only if $x_1 - x_2 \in A^*$ and $x_1 \in D$. Use Lemma 1 to find

$$\begin{aligned} \mathrm{E}\{c_{12}\} = \mathrm{P}[c_{12} = 1] \leqslant \mathrm{P}[x_1 - x_2 \in A^*, x_1 \in D] \\ &= \mathrm{P}[x_1 - x_2 \in A^*]\ \mathrm{P}[x_1 \in D] \\ &= \lambda(1-(1-2r)^k) \\ &\leqslant 2k\lambda r. \end{aligned}$$

Since

$$\operatorname{var} c_{12} = \mathrm{E}\{c_{12}(1-\mathrm{E}\{c_{12}\})\},$$

(a) is proven.

(b) $\operatorname{cov}(c_{12}, c_{13}) = \mathrm{E}\{c_{12}c_{13}\} - \mathrm{E}\{c_{12}\}\ \mathrm{E}\{c_{13}\}$. By (a), the last term is of order $r^2\lambda^2$, therefore consider $\mathrm{E}\{c_{12}\, c_{13}\}$ only. Proceeding analogously to (a), $\mathrm{P}[c_{12} = 1, c_{13} = 1] \leqslant \mathrm{P}[x_1 - x_2 \in A^*]\,\mathrm{P}[x_1 - x_3 \in A^*]\,\mathrm{P}[x_1 \in D] \leqslant \lambda^2\, 2kr$.

2. The main theorem

In the above, A and r were allowed to vary independently of n. Now take $A = A_n$, $r = r_n$, and consider $s_n = s_n(A_n)$ and $t_n = t_n(A_n)$. The following theorem gives conditions on the sequence (A_n) to get limit results.

Theorem 2.1. (a) *Let $\{r_n\}_{n=1}^{\infty}$ be a sequence of real numbers, $0 < r_n < \frac{1}{2}$, let $\{A_n\}_{n=1}^{\infty}$ be subsets of $\mathbf{R}^k$, with $A_n \subset B(r_n)$, such that for $n \to \infty$,*

(i) $n^2\ \lambda(A_n) \to \infty$;

(ii) $r_n \to 0$;

(iii) $nr_n\lambda(A_n) \to 0$.

Then $t_n(A_n)$ is asymptotically normal.

(b) *Let $\{A_n\}_{n=1}^{\infty}$ be subsets of C^k such that for $n \to \infty$,*

(i) $n^2\ \lambda(A_n) \to \infty$;

(ii) $\lambda(A_n) \to 0$.

Then $s_n(A_n)$ is asymptotically normal.

Proof. From Bernstein's Lemma [3, Theorem 8.3] one immediately concludes that (1.3) and conditions (a) (ii), (iii) are sufficient to infer asymptotic normality for t_n from asymptotic normality for s_n. Since further conditions (a) (i)–(iii) imply (b) (i), (ii) it is clear that (a) is proven whenever (b) is. Part (b) will be proven by showing that all moments of the standardized s_n tend to the moments of a standard normal distribution. That this is sufficient for asymptotic normality is known as the Second Limit Theorem [3, Theorem 6.14].

For any r.v. a write $a' = a - \mathrm{E}\{a\}$ for the centered r.v. Further put $\sigma_n^2 = \operatorname{var} s_n$. For fixed integers r, n with $1 \leqslant r \leqslant n$, consider

$$\mu_{rn} = \mathrm{E}\{(s_n'/\sigma_n)^r\} = \mathrm{E}\left\{\left(\sum b_{ij}'\right)^r\right\}/\sigma_n^r.$$

Working out the r^{th} power in the numerator and substituting the order of σ_n^r from (1.2), one finds

$$\mu_{rn} \sim (\tfrac{1}{2}n^2\lambda)^{-r/2} \sum_\tau \pi(\tau), \tag{2.1}$$

where

$$\pi(\tau) = \mathrm{E}\{b'_{i_1 j_1} b'_{i_2 j_2} \ldots b'_{i_r j_r}\}, \tag{2.2}$$

and $\tau = (i_1, j_1; i_2, j_2; \ldots; i_r, j_r)$ ranges over T, the set of all ordered r-tuples of pairs (i, j) with $1 \leqslant i < j \leqslant n$. Since the x_i are i.i.d., any permutation of $\{1, 2, \ldots, n\}$, applied to the indices in τ will not alter the value of $\pi(\tau)$. Hence one can collect equal terms $\pi(\tau)$ and find

$$\sum_{\tau \in T} \pi(\tau) = \sum_{\tau \in T^*} \nu(\tau)\, \pi(\tau), \tag{2.3}$$

where T^* is T modulo permutations of $\{1, 2, \ldots, n\}$, and $\nu(\tau)$ is the number of elements of T in modulo-class τ. In factorization (2.3), the number of terms depends only on r, each $\nu(\tau)$ depends on n only, and $\pi(\tau)$ only on λ.

It will be convenient to think of $\tau \in T^*$ as an undirected graph with vertex set $\{i_1, j_1, i_2, j_2, \ldots, i_r, j_r\}$ and edges (i_k, j_k), $k = 1, 2, \ldots, r$. Note that τ has r edges, some of which may coincide, and that the number of vertices $l(\tau)$ is the number of different indices in τ. Define $c(\tau)$ to be the number of components of τ.

One is interested in those $\tau \in T^*$ for which $\nu(\tau)\, \pi(\tau)$ is of order $(n^2\lambda)^{r/2}$ or higher. It will be shown that for all except one $\tau \in T^*$ the contribution $\nu(\tau)\, \pi(\tau)$ to the sum (2.3) is in fact of lower order. This will follow from a lemma, whose proof is deferred until Section 3.

Lemma 2.2. (a) $\nu(\tau) = \mathrm{O}(n^{l(\tau)})$, $n \to \infty$.
(b) *If τ has a singleton index,* $\pi(\tau) = 0$.
(c) $\pi(\tau) = \mathrm{O}(\lambda^{l(\tau) - c(\tau)})$ *for* $\lambda \downarrow 0$.

Consider the right-hand side of (2.3). Since any τ with more than $\frac{1}{2}r$ different indices must have at least one singleton index, by Lemma 2(b), such τ do not contribute to (2.3). Suppose therefore $l(\tau) \leq \frac{1}{2}r$. Substituting the results of Lemma 2, one obtains

$$\nu(\tau)\ \pi(\tau)/\sigma_n^r = \mathrm{O}((n^2\ \lambda)^{(l(\tau)-r)/2}\ \lambda^{l(\tau)/2-c(\tau)}). \tag{2.4}$$

Since a component has at least two vertices, one has $\frac{1}{2}l(\tau) \geq c(\tau)$. Therefore (2.4) $\to 0$ unless $l(\tau) = r = 2c(\tau)$. This implies that

$$\mu_{rn} \to 0$$

for odd r, and

$$\mu_{rn} = \nu(\tau_0)\ \pi(\tau_0)/\tfrac{1}{2}n^2\lambda)^{r/2}$$

for even r, where $\tau_0 = (1,2;\ 1,2;\ 3,4;\ 3,4;\ \ldots;\ \frac{1}{2}(r-2), \frac{1}{2}r;\ \frac{1}{2}(r-2), \frac{1}{2}r)$. One finds

$$\pi(\tau_0) = \mathrm{E}\{(b'_{12})^{r/2}\} = \lambda^{r/2}\ (1-\lambda)^{r/2}.$$

Further $\nu(\tau_0) = \binom{n}{r}a_0$, where a_0 is the number of terms in $(\Sigma_{1\leq i<j\leq r} b_{ij})^r$, consisting of a product of $\frac{1}{2}r$ squares with $\frac{1}{2}r$ pairs of distinct indices. Some combinatorics lead to

$$a_0 = (r!)^2\ 2^{-r}/(\tfrac{1}{2}r)!,$$

so that

$$\nu(\tau_0) = n^r\ r!\ 2^{-r}/(\tfrac{1}{2}r)! + \mathrm{O}(n^{r-1}), \quad n \to \infty.$$

Hence for even r one finds

$$\mu_{rn} \to r!/(\tfrac{1}{2}r)!\ 2^{r/2},$$

the r^{th} moment of a standard normal distribution. This completes the proof. □

3. Proof of Lemma 2.2

(a) $\nu(\tau) = \mathrm{O}(n^{l(\tau)})$, $n \to \infty$. First suppose that the $l(\tau)$ different indices are fixed, $\{1, 2, \ldots, l(\tau)\}$ say. There exists some number m of terms in modulo-class $\tau \in T^*$ that have indices in exactly the set $\{1, 2, \ldots, l(\tau)\}$. For any choice of $l(\tau)$ different indices from $\{1, 2, \ldots, n\}$ there are m terms in $\tau \in T^*$. Hence the total number of terms is

$$\nu(\tau) = \binom{n}{l(\tau)} \cdot m = \mathrm{O}(n^{l(\tau)}), \quad n \to \infty.$$

(b) If τ has a singleton index, $\pi(\tau) = 0$. A more general statement will be proven: if the indices in τ can be partitioned into two disjoint sets such that exactly one pair (i_k, j_k) has its elements in different sets, then $\pi(\tau) = 0$. Viewed as a graph, τ has an isthmus or a pendant. Let

$$\{i_1, j_1, i_2, j_2, ..., i_{k-1}, j_{k-1}, i_k\}, \qquad \{j_k, i_{k+1}, j_{k+1}, ..., i_r, j_r\}$$

be disjoint sets. If τ has a singleton index, $k = 1$. Since $\pi(\tau)$ is independent of the origin in C^k, one may choose the origin to coincide with x_{j_k}. Under this condition the two sets of r.v.'s

$$\{b_{i_h j_h} \mid h = 1, 2, ..., k\}, \qquad \{b_{i_h j_h} \mid h = k + 1, k + 2, ..., r\}$$

are independent, so that

$$\pi(\tau) = \mathrm{E}\{b'_{i_1 j_1} \cdots b'_{i_k j_k} \mid x_{j_k} = 0\}\, \mathrm{E}\{b'_{i_{k+1} j_{k+1}} \cdots b'_{i_r j_r} \mid x_{j_k} = 0\}.$$

Now in the first factor one may forget the conditioning on x_{j_k} and instead condition on $x_{i_k} = 0$. Then $b_{i_k j_k}$ is independent of the other r.v.'s and one can factor again. Since

$$\mathrm{E}\{b'_{i_k j_k} \mid x_{i_k} = 0\} = 0,$$

one finds $\pi(\tau) = 0$.

Implicitly the above reasoning shows that $\mathrm{E}\{b_{ij}\}$ can be factored out of $\mathrm{E}\{b_{i_1 j_1} \dots b_{i_r j_r}\}$ whenever b_{ij} represents an isthmus or pendant in the graph τ.

(c) $\pi(\tau) = \mathrm{O}(\lambda^{l(\tau)-c(\tau)})$, $\lambda \downarrow 0$. Assume that the statement has been proven for all terms of the form (2.2) with the number of factors b'_{ij} less than r. Now consider $\pi(\tau)$ with $\tau = (i_1, j_1; ...; i_r, j_r)$. Then

$$\pi(\tau) = \mathrm{E}\left\{\prod_{k=1}^{r} (b_{i_k j_k} - \lambda)\right\} = \mathrm{E}\{\Pi b_{i_k j_k}\} - \lambda \sum_{k=1}^{r} \mathrm{E}\left\{\prod_{h \neq k} b'_{i_h j_h}\right\}. \tag{3.1}$$

First let t_τ be a subset of the pairs of τ such that t_τ forms a maximal tree in τ:

$$l(t_\tau) = l(\tau), \qquad c(t_\tau) = c(\tau),$$

and every edge of t_τ is isthmus or pendant. Then the number of pairs (edges) in t_τ equals $l(\tau) - c(\tau)$ [1, ch. 16]. Using the fact that the b_{ij} are 0–1 variables, and the remark at the end of the proof of part (b), one finds

$$\mathrm{E}\left\{\prod_{(i_k, j_k) \in \tau} b_{i_k j_k}\right\} \leqslant \mathrm{E}\left\{\prod_{(i_k, j_k) \in t_\tau} b_{i_k j_k}\right\} = \lambda^{l(\tau)-c(\tau)}.$$

Hence the first term in (3.1) is $O(\lambda^{l(\tau)-c(\tau)})$, the order stated. Next consider the second term in (3.1). Note that the expectations in the summation are of the form $\pi(\delta_k)$, where δ_k stands for the graph τ with the k^{th} edge deleted. Hence, by assumption

$$\pi(\delta_k) = O(\lambda^{l(\delta_k)-c(\delta_k)}),$$

so that the exponent of λ in the order of the second term in (3.1) is

$$\max_k\,[l(\delta_k) - c(\delta_k) + 1].$$

It is easily checked that

$$l(\delta) - c(\delta) + 1 \geqslant l(\tau) - c(\tau)$$

for a graph δ differing from τ by exactly one edge, only if the deleted edge is an isthmus or pendant in τ. But in that case $\pi(\tau) = 0$ by (b). Hence the second term in (3.1) is of lower order.

4. A multivariate extension

The next theorem is a multivariate analog of Theorem 2.1. Its proof proceeds along the same lines as in the univariate case.

Theorem 4.1. (a) *Let $\{r_n\}_{n=1}^{\infty}$ be a sequence of real numbers $0 < r_n < \frac{1}{2}$ and for $i = 1, 2, \ldots, m$ let $\{A_{ni}\}_{n=1}^{\infty}$ be a sequence of subsets of $\mathbf{R}^k$ such that for $n \geqslant 1$*

$$\bigcup_{i=1}^{m} A_{ni} \subset B(r_n),$$

and $A_{n1}, A_{n2}, \ldots, A_{nm}$ are disjoint. Put $t_{ni} = t_n(A_{ni})$ for $i = 1, 2, \ldots, m$ and $n \geqslant 1$. As $n \to \infty$, let

(i) $n^2\,\lambda(A_{ni}) \to \infty$ $(1 \leqslant i \leqslant m)$,

(ii) $r_n \to 0$,

(iii) $n\,r_n\,\lambda(A_{ni}) \to 0$ $(1 \leqslant i \leqslant m)$.

Then $(t_{n1}, t_{n2}, \ldots, t_{nm})$ have simultaneously an asymptotically normal distribution and are asymptotically independent.

(b) *Let $\{A_{ni}\}_{n=1}^{\infty}$ be a sequence of subsets of C^k $(1 \leqslant i \leqslant m)$ such that for $n \geqslant 1$, $A_{n1}, A_{n2}, \ldots, A_{nm}$ are disjoint. Put $s_{ni} = s_n(A_{ni})$ for $1 \leqslant i \leqslant m$ and $n \geqslant 1$. As $n \to \infty$, let*

(i) $n^2\,\lambda(A_{ni}) \to \infty$ $(1 \leqslant i \leqslant m)$,

(ii) $\lambda(A_{ni}) \to 0$ $(1 \leqslant i \leqslant m)$.

Then $(s_{n1}, s_{n2}, ..., s_{nm})$ *have simultaneously an asymptotically normal distribution and are asymptotically independent.*

Proof. First note that similarly to the univariate case, condition (a) (iii) implies $\mathrm{var}(s_{ni}-t_{ni})/\mathrm{var}\, s_{ni} \to 0$ for all i. From the converging together theorem [2, Theorem 4.2] one then concludes that part (a) is proven whenever part (b) is. Assume (b) (i), (ii) hold. Abbreviate

$$\sigma_{ni}^2 = \mathrm{var}\, s_{ni}, \qquad \lambda_i = \lambda(A_{ni}),$$

and for fixed integers $r_1, r_2, ..., r_m \geq 0$ and $n > r = r_1 + r_2 + ... + r_m$ consider

$$\mu_{r_1 r_2 ... r_m; n} = \mathrm{E}\left\{\prod_{i=1}^{n} (s'_{ni}/\sigma_{ni})^{r_i}\right\}.$$

The denominator is known:

$$\sigma_{ni}^2 = \tfrac{1}{2} n(n-1)\, \lambda_i(1-\lambda_i).$$

In the numerator substitute

$$s'_{nl} = \sum_{i<j} b'_{ij;l}$$

where $b_{ij;l} = 1$ if $x_i - x_j \in A_{nl}$ and 0 otherwise. Then as in the univariate case one obtains

$$\mu_{r_1 r_2 ... r_m; n} \sim \left\{(\tfrac{1}{2} n^2)^{r/2} \prod_{i=1}^{m} \lambda_i^{r_i}\right\}^{-1} \sum_{\tau} \nu(\tau)\, \pi(\tau), \tag{4.1}$$

where now $\pi(\tau)$ is the expectation of a product of factors $b'_{ij;l}$. In this situation τ may be thought of as a graph with m different kinds of edges, say white, red, ..., black edges, the color of each edge determined by the last subscript in the corresponding $b_{ij;l}$.

Checking the proof of Lemma 2.2, it is easily seen that (a) and (b) remain valid in this more general setting, and that part c holds if one puts $\lambda = \max \lambda_i$. Then as before one concludes that only graphs of the form τ_0 contribute to (4.1). Different colorings must now be considered, however. First consider the case where $\pi(\tau)$ does not contain any products of the form $b_{ij;l}\, b_{ij;k}$, $k \neq l$. Then all r_i are even and $\pi(\tau) \sim \Pi \lambda_i^{r_i/2}$, as is easily seen from the independence of the pairs of factors in $\pi(\tau)$. Further,

$$\nu(\tau) \sim \prod_{i=1}^{m} \frac{n^{r_i}\, r_i!\, 2^{-r_i}}{(\frac{1}{2} r_i)!}.$$

If $\pi(\tau)$ does contain a mixed term $b_{ij;l}\, b_{ij;k}$, the order of $\pi(\tau)$ is strictly greater than $\Pi\lambda_i^{r_i/2}$, so that those contributions vanish in the limit. One therefore finds that

$$\mu_{r_1 r_2 \dots r_m;n} \to \begin{cases} 0 & \text{if not all } r_i \text{ are even.} \\ \prod \dfrac{r_i!}{2^{r_i/2}\,(\frac{1}{2}r_i)!} & \text{otherwise.} \end{cases}$$

Hence all moments of the joint distribution of $(s_{n1}, s_{n2}, \dots, s_{nm})$ tend to the corresponding moments of a multivariate normal vector, with independent components.

Acknowledgment

I am grateful to A.A. Balkema for helpful comments.

References

[1] C. Berge, The Theory of Graphs and its Applications (Methuen, London, 1960).
[2] P. Billingsley, Convergence of Probability Measures (Wiley, New York, 1968).
[3] P.A.P. Moran, Introduction to Probability Theory (Clarendon, Oxford, 1968).